

CT/EP 99/03469
BUNDESREPUBLIK DEUTSCHLAND



49041
EP 99/3469

REC'D 29 JUN 1999

WIPO PCT

PRIORITY
DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

Bescheinigung

Die BASF Aktiengesellschaft in Ludwigshafen/Deutschland hat
eine Patentanmeldung unter der Bezeichnung

"Neue Peptidfragmente zur Reinigung von Proteinen"

am 20. Mai 1998 beim Deutschen Patent- und Markenamt
eingereicht.

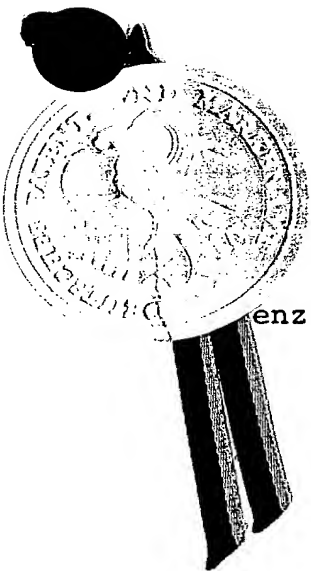
Die angehefteten Stücke sind eine richtige und genaue
Wiedergabe der ursprünglichen Unterlagen dieser Patent-
anmeldung.

Die Anmeldung hat im Deutschen Patent- und Markenamt vor-
läufig die Symbole C 07 K, C 12 N und C 07 H der Interna-
tionalen Patentklassifikation erhalten.

München, den 28. Dezember 1998
Deutsches Patent- und Markenamt

Der Präsident

Im Auftrag



Zeichen: 198 22 823.6

Joost

M 21.03.98

Patentansprüche

1. Peptidfragment mit der allgemeinen Sequenz

5

His-X¹-His-X²-X³-X⁴-Cys-X⁵-X⁶-Cys,

wobei die Variablen X¹ bis X⁶ in der Sequenz die folgende Bedeutung haben:

10

X¹ = eine Aminosäure ausgewählt aus der Gruppe Ala, Val, Phe, Ser, Met, Trp, Tyr, Asn, Asp oder Lys und die Variablen X² bis X⁶ eine Aminosäure ausgewählt aus der Gruppe Gly, Ala, Val, Leu, Ile, Phe, Pro, Ser, Thr, Cys, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His oder

15

X² = eine Aminosäure ausgewählt aus der Gruppe Val, Ile, Phe, Pro, Trp, Tyr, Gln, Glu oder Arg und die Variablen X¹, X³ bis X⁶ eine Aminosäure ausgewählt aus der Gruppe Gly, Ala, Val, Leu, Ile, Phe, Pro, Ser, Thr, Cys, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His oder

20

X³ = eine Aminosäure ausgewählt aus der Gruppe Gly, Ile, Thr, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His und die Variablen X¹, X², X⁴ bis X⁶ eine Aminosäure ausgewählt aus der Gruppe Gly, Ala, Val, Leu, Ile, Phe, Pro, Ser, Thr, Cys, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His oder

25

X⁴ = eine Aminosäure ausgewählt aus der Gruppe Val, Phe, Pro, Cys, Met, Trp, Asn, Glu, Arg oder His und die Variablen X¹ bis X³, X⁵, X⁶ eine Aminosäure ausgewählt aus der Gruppe Gly, Ala, Val, Leu, Ile, Phe, Pro, Ser, Thr, Cys, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His oder

30

35

X⁵ = eine Aminosäure ausgewählt aus der Gruppe Gly, Ser, Cys, Met, Trp, Asn, Glu, Lys oder Arg und die Variablen X¹ bis X⁴, X⁶ eine Aminosäure ausgewählt aus der Gruppe Gly, Ala, Val, Leu, Ile, Phe, Pro, Ser, Thr, Cys, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His oder

40

H 21.05.99
2

X⁶ = eine Aminosäure ausgewählt aus der Gruppe Phe, Pro, Ser, Cys, Trp, Tyr oder Gln und die Variablen X¹ bis X⁵ eine Aminosäure ausgewählt aus der Gruppe Gly, Ala, Val, Leu, Ile, Phe, Pro, Ser, Thr, Cys, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His und

wobei mindestens eine der Variablen X¹ bis X⁶ in der Sequenz unabhängig voneinander Gln oder Asn bedeutet

10 2. Peptidfragment nach Anspruch 1, in der die Variablen X¹ bis X⁶ die gemäß Anspruch 1 genannte Bedeutung haben, wobei mindestens eine der Variablen X¹ bis X⁶ in der Sequenz unabhängig voneinander Lys oder Arg bedeutet.

15 3. Peptidfragment nach Anspruch 1 oder 2, in der die Variablen X¹ bis X⁶ in der Sequenz unabhängig voneinander folgende Bedeutung haben:

20 X¹ = eine Aminosäure ausgewählt aus der Gruppe Ala, Val, Phe, Ser, Met, Trp, Tyr, Asn, Asp oder Lys;

X² = eine Aminosäure ausgewählt aus der Gruppe Val, Ile, Phe, Pro, Trp, Tyr, Gln, Glu oder Arg;

25 X³ = eine Aminosäure ausgewählt aus der Gruppe Gly, Ile, Thr, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg oder His;

X⁴ = eine Aminosäure ausgewählt aus der Gruppe Val, Phe, Pro, Cys, Met, Trp, Asn, Glu, Arg oder His;

30 X⁵ = eine Aminosäure ausgewählt aus der Gruppe Gly, Ser, Cys, Met, Trp, Asn, Glu, Lys oder Arg;

35 X⁶ = eine Aminosäure ausgewählt aus der Gruppe Phe, Pro, Ser, Cys, Trp, Tyr oder Gln.

4. Peptidfragment nach den Ansprüchen 1 bis 3, in der die Variablen X¹ bis X⁶ in der Sequenz unabhängig voneinander folgende Bedeutung haben:

40 X¹ = eine Aminosäure ausgewählt aus der Gruppe Phe, Ser, Asn, Asp oder Lys;

45 X² = eine Aminosäure ausgewählt aus der Gruppe Val, Ile, Phe, Pro, Gln, Glu oder Arg;

3. 12105-000

x⁴ = eine Aminosäure ausgewählt aus der Gruppe Val, Phe, Cys, Met, Trp, Asn, Arg oder His;

10 x^6 = eine Aminosäure ausgewählt aus der Gruppe Phe, Ser, Cys,
oder Tyr.

15

20

X³ = Gly, Thr oder Tyr;

25

X⁵ = Gly oder Lys;

30

His-Gln-His-Glu-Gly-Arg-Cys-Lys-Glu-Cys

His-Asn-His-Arg-Tyr-Gly-Cys-Gly-Cys-Cys

35 His-Arg-His-Gly-Thr-Asn-Cys-Leu-Lys-Cys

His-Ile-His-Gln-Ser-Asn-Cys-Gln-Val-Cys.

7. Fusionsprotein enthaltend ein Proteinfsegment gemäß den
Ansprüchen 1 bis 6.

8. Nukleinsäurefragment kodierend für ein Proteinfsegment gemäß den Ansprüchen 1 bis 6.

45 9. Nukleinsäuren enthaltend ein Nukleinsäurefragment gemäß
Anspruch 8.

14.05.99
4

10. Nukleinsäuren kodierend für ein Fusionsprotein gemäß Anspruch 7.
11. Vektor enthaltend ein Nukleinsäurefragment gemäß Anspruch 8 oder 10.
12. Verfahren zur Herstellung von Fusionsproteinen gemäß Anspruch 7, dadurch gekennzeichnet, daß man ein Nukleinsäurefragment gemäß Anspruch 8 mit einem Gen, das für ein Protein kodiert, fusioniert.
13. Verfahren zur Reinigung von Fusionsproteinen gemäß Anspruch 7, dadurch gekennzeichnet, daß man
- a) Flüssigkeiten, die das Fusionsprotein enthalten, so mit immobilisierten Metallionen in Kontakt bringt, daß sich zwischen den Metallionen und dem Fusionsprotein eine Affinitätsbindung ausbilden kann,
 - b) Entfernen nicht gebundener in der Flüssigkeit enthaltener Substanzen,
 - c) Eluieren des gebundenen Fusionsproteins, in dem die Affinitätsbindung durch Änderung des Flüssigkeits milieus aufgehoben wird und
 - d) Sammlung des gereinigten Fusionsprotein.
14. Verwendung eines Proteinfragments gemäß den Ansprüchen 1 bis 6 oder eines Nukleinsäurefragments gemäß Anspruch 8 zur Reinigung von Proteinen .
15. Verfahren zur Herstellung von Proteinfragmenten, die in der Lage sind mit immobilisierten Metallionen eine reversible Affinitätsbindung einzugehen, dadurch gekennzeichnet, daß man folgende Schritte durchführt:
- a) Herstellung einer Nukleinsäurebibliothek ausgehend von einer beliebigen Nukleinsäuresequenz, die für ein Proteinfragment der Sequenz
 $\text{His-X}^1\text{-His-X}^2\text{-X}^3\text{-X}^4\text{-Cys-X}^5\text{-X}^6\text{-Cys,}$
kodiert, wobei die Histidin- und Cysteinreste der Sequenz in der Nukleinsäurebibliothek konserviert werden,

H 2 1 . 1 5 . 0 0 0
5

- b) Fusion der Nukleinsäuren der Bibliothek an ein Reporter-
gen, das die Detektion des durch die erhaltene Nuklein-
säure kodierten Fusionsproteins über seine Bindung an die
immobilisierten Metallionen ermöglicht und

5

- c) Auswahl der Nukleinsäuresequenzen, die eine gegenüber der
in der natürlichen *Helicobacter pilori* ATPase-439 Sequenz
um mindestens 1,5 fach stärkere reversible Bindung an die
immobilisierten Metallionen aufweisen.

10

16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, das als
Reporter-Gen das egf-Protein von *Aequoria victoria* verwendet
wird.

15

17. Verfahren zur Detektion von Proteinen, dadurch gekennzeich-
net, daß man aus einer Proteinmischung einzelne Proteine, die
ein Proteinfragment gemäß Anspruch 1 enthalten, über Anti-
körper nachweist, die gegen das Proteinfragment gerichtet
sind.

20

18. Verwendung eines Proteinfragments gemäß den Ansprüchen 1 bis
6 oder eines Nukleinsäurefragments gemäß Anspruch 8 zur
Reinigung von Proteinen.

25

30

35

40

45

11.21.05.99

Neue Peptidfragmente zur Reinigung von Proteinen

Beschreibung

- 5 Die Erfindung betrifft neue Peptidfragmente, Fusionsproteine enthaltend die Peptidfragmente, Verfahren zur ihrer Herstellung sowie die Verwendung der Peptidfragmente. Die Erfindung betrifft auch ein Verfahren zur Reinigung der Fusionsproteine und ein
- 10 Verfahren zur Detektion von Proteinen.

Weiterhin betrifft die Erfindung Nukleinsäuren, die für die Peptidfragmente oder für die Fusionsproteine kodieren und Vektoren, die die Nukleinsäuren enthalten.

- 15 Durch die moderne Molekularbiologie ist es möglich geworden fast beliebige Proteine, Peptide oder deren Derivate in nahezu unbegrenzten Mengen herzustellen. Die Proteinreinigung erweist sich dabei häufig als der limitierende und nicht selten ineffiziente
- 20 und damit letztlich kostenbestimmende Faktor.

- Es wurden deshalb eine Reihe von Techniken zur Reinigung von Proteinen entwickelt. Für die Reinigung von Proteinen aus Zellüberständen, Zellrohextrakten oder Zellen werden üblicherweise
- 25 Techniken wie Salzpräzipitation oder Präzipitation mit organischen Lösungsmitteln, Ultrafiltration, Dialyse, Gelelektrophorese, Isoelektrische Fokussierung, Chromatofokussierung, Ionenaustauscher- oder Gelchromatographie, Hydrophobe Chromatographie, Immunopräzipitation oder IMAC (= Immobilized Metal Affinity Chromatographie) verwendet. Es können einzelne Methoden zur
- 30 Reinigung verwendet werden, in der Regel ist aber eine Kombination verschiedener Techniken erforderlich. Eine Übersicht über übliche Reinigungsmethoden ist beispielsweise den Lehrbüchern Protein Purification Process Engineering (Ed. R.G. Harrison, 1994, Marcel Dekker, Inc., New York, page 209 - 316, ISBN
- 35 0-8247-9009-X), Protein Purification (Ed. R.K. Scopes, 1994, Springer Verlag New York, chapter 4 - 7, ISBN 0-387-94072-3) und Methods for Protein Analysis (Ed. R.A. Copeland, 1994 Chapman & Hall, page 59 - 112, ISBN 0-412-03741-6) zu entnehmen. Für die Proteinreinigung ist es wichtig, daß die Reinigung so schonend
- 40 wie möglich, so selektiv wie möglich und möglichst rasch erfolgt. Dabei sollen die Proteinverluste so gering wie irgend möglich gehalten werden. Viele der Proteinreinigungsmethoden sind nicht ausreichend selektiv, für nur geringe Proteinmengen geeignet und/oder sehr teuer.

M I L I O N E N

Die von Porath et al. (Nature, Vol.258, 1975: 598 - 599)

beschriebene Methode zur Proteinreinigung mit Hilfe der IMAC
(= Immobilized Metal Affinity Chromatographie) ist ein guter
Kompromiss zwischen den genannten Anforderungen an eine optimale
5 Reinigung. Diese Methode hat jedoch noch einige Nachteile. So
binden beispielsweise nicht alle Metallionen gleich gut an das
Trägermaterial, so daß die Ionen zum Teil ausgewaschen werden und
so das Wertprodukt verunreinigen. Viele Proteine binden gar nicht
an das Chromatographiematerial und lassen sich so nicht reinigen
10 oder binden zu schwach, so daß sie bei den erforderlichen Wasch-
schritten des Säulenmaterials schon eluiert werden. Dies führt
zu unerwünschten Produktverlusten. Da die Selektivität in der
Regel für eine "Ein-Schritt-Reinigung" nicht ausreichend ist, im
Gegensatz zur Reinigung über biospezifische Affinitätsreinigungs-
15 methoden wie beispielsweise eine Reinigung über Antikörper, sind
weitere Reinigungsschritte erforderlich.

Um eine breitere Palette an Proteinen mit dieser Methode reinigen
zu können, wurde verschieden sogenannte Tags wie Polyhistidin,
20 His-Trp, His-Tyr oder (His-Asp)_n entwickelt (siehe Sporeno et al.,
J. Biol. Chem. 269 (15), 1994: 10991 - 10995, Le Grice et al.,
Eur. J. Biochem., 187 (2), 1990: 307 - 314, Reece et al., Gene,
126 (1), 1993: 105 - 107, De Vos et al., Nucl. Acid. Res., 22
(7), 1994: 1161 - 1166, Feng et al., J. Biol. Chem. 269 (3),
25 1994: 2342 - 2348, Hochuli et al., Biotechnology, 1988: 1321 -
1325, Patwardhan et al., J. Chromatography A, 787, 1997: 91 -
100, Hutchens et al., J. Chromatography, 604, 1992: 133 - 141).
Diese Tags werden mit den zu reinigenden Protein mit Hilfe der
Molekularbiologie auf Nukleinsäureebene verbunden. Durch diese
30 Tags konnte die Proteinreinigung in einigen Bereichen weiter ver-
bessert werden. Jedoch auch diese Methode hat nach wie vor einige
Nachteile. Nach wie vor läßt sich nicht sicher Vorhersagen, ob
ein Protein über diese Methode aufgereinigt werden kann (siehe
Immobilized Metal Ion Affinity Chromatography, L. Kågedal, page
227 - 251 in Protein Purification, Eds. J.C. Janson, L. Rydén,
35 1989, VCH Publishers, Inc., New York, ISBN 0-89573-122-3), das
heißt auch diese Methode ist nicht auf jedes Protein anwendbar.
Auch hier können Metallionen ausgewaschen werden oder Proteine so
schwach binden, daß sie während der Waschschrte zum Teil ver-
loren gehen. Auch die Selektivität läßt noch zu wünschen übrig.
40 Außerdem ist die Beladungskapazität des Säulenmaterials mit den
zu reinigenden Proteinen teilweise noch zu gering, so daß für
eine Reinigung viel Säulenmaterial verwendet werden muß. Auch
die Proteinausbeute ist noch nicht ausreichend. Dies führt zu
unnötigen Kosten.

M 21.05.99
3

Von Volz et al. wurde die Reinigung zweier ATPasen von *Helicobacter pylori* ohne die Verwendung einer zusätzlichen His-Tag-Sequenz beschrieben. Diese ATPasen enthalten natürliche Metallbindungsstellen, die eine Reinigung über IMAC ermöglichen. Eigene Arbeiten zeigten, daß diese Bindungsstellen für eine effiziente Aufreinigung beliebiger Proteine eine zu geringe Bindungsaffinität aufweisen.

Es bestand daher die Aufgabe weitere Tags für die Proteinreinigung mittels IMAC zur Verfügung zu stellen, die die oben genannten Nachteile nicht besitzen und dadurch beispielsweise eine breitere Anwendung der Tags und/oder eine höhere Beladungsdichte der des Säulenmaterials ermöglichen. Die eine höhere Selektivität zeigen und so die Reinigung vereinfachen.

Die Aufgabe wurde durch die erfindungsgemäßen Peptidfragmente mit der allgemeinen Sequenz

His-X¹-His-X²-X³-X⁴-Cys-X⁵-X⁶-Cys gelöst,

wobei die Variablen X¹ bis X⁶ in der Sequenz die folgende Bedeutung haben:

X¹ = eine Aminosäure ausgewählt aus der Gruppe Ala, Val, Phe, Ser, Met, Trp, Tyr, Asn, Asp oder Lys und die Variablen X² bis X⁶ eine Aminosäure ausgewählt aus der Gruppe Gly, Ala, Val, Leu, Ile, Phe, Pro, Ser, Thr, Cys, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His oder

X² = eine Aminosäure ausgewählt aus der Gruppe Val, Ile, Phe, Pro, Trp, Tyr, Gln, Glu oder Arg und die Variablen X¹, X³ bis X⁶ eine Aminosäure ausgewählt aus der Gruppe Gly, Ala, Val, Leu, Ile, Phe, Pro, Ser, Thr, Cys, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His oder

X³ = eine Aminosäure ausgewählt aus der Gruppe Gly, Ile, Thr, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His und die Variablen X¹, X², X⁴ bis X⁶ eine Aminosäure ausgewählt aus der Gruppe Gly, Ala, Val, Leu, Ile, Phe, Pro, Ser, Thr, Cys, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His oder

X⁴ = eine Aminosäure ausgewählt aus der Gruppe Val, Phe, Pro, Cys, Met, Trp, Asn, Glu, Arg oder His und die Variablen X¹ bis X³, X⁵, X⁶ eine Aminosäure ausgewählt aus der Gruppe Gly, Ala, Val, Leu, Ile, Phe, Pro, Ser, Thr, Cys, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His oder

4

x^5 = eine Aminosäure ausgewählt aus der Gruppe Gly, Ser, Cys, Met, Trp, Asn, Glu, Lys oder Arg und die Variablen x^1 bis x^4 , x^6 eine Aminosäure ausgewählt aus der Gruppe Gly, Ala, Val, Leu, Ile, Phe, Pro, Ser, Thr, Cys, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His oder

x^6 = eine Aminosäure ausgewählt aus der Gruppe Phe, Pro, Ser, Cys, Trp, Tyr oder Gln und die Variablen x^1 bis x^5 eine Aminosäure ausgewählt aus der Gruppe Gly, Ala, Val, Leu, Ile, Phe, Pro, Ser, Thr, Cys, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg, His und

wobei mindestens eine der Variablen x^1 bis x^6 in der Sequenz unabhängig voneinander Gln oder Asn bedeutet.

Vorteilhafterweise bedeutet in der Sequenz mindestens eine der Variablen x^1 bis x^6 außerdem unabhängig voneinander Lys oder Arg. Weitere vorteilhafte in den Variablen x^1 bis x^6 enthaltene Aminosäuren sind Glu, Lys, Arg, Tyr, Cys, Lys, His, Asp oder Met. Bevorzugt sind die Aminosäuren Cys, Glu, Lys, Tyr oder Arg enthalten, besonders bevorzugt Cys. Diese Aminosäuren tragen zu einer vorteilhaften Bindung der Peptidfragmente an die immobilisierten Metallionen bei. Außerdem sind vorteilhafterweise nicht mehr als vier bevorzugt drei Histidinreste in Folge in der Sequenz enthalten.

Weiter haben die Variablen x^1 bis x^6 in der Sequenz unabhängig voneinander folgende vorteilhafte bevorzugte Bedeutung:

x^1 = eine Aminosäure ausgewählt aus der Gruppe Ala, Val, Phe, Ser, Met, Trp, Tyr, Asn, Asp oder Lys, besonders bevorzugt Phe, Ser, Asn, Asp oder Lys, ganz besonders bevorzugt Asn;

x^2 = eine Aminosäure ausgewählt aus der Gruppe Val, Ile, Phe, Pro, Trp, Tyr, Gln, Glu oder Arg, besonders bevorzugt Val, Ile, Phe, Pro, Gln, Glu oder Arg, ganz besonders bevorzugt Gln, Glu oder Arg;

x^3 = eine Aminosäure ausgewählt aus der Gruppe Gly, Ile, Thr, Met, Trp, Tyr, Asn, Gln, Asp, Glu, Lys, Arg oder His, besonders bevorzugt Gly, Ile, Thr, Met, Trp, Tyr, Asn, Asp, Glu, Arg oder His, ganz besonders bevorzugt Gly, Thr oder Tyr;

5 2 1 0 5 0 0

x⁴ = eine Aminosäure ausgewählt aus der Gruppe Val, Phe, Pro, Cys, Met, Trp, Asn, Glu, Arg oder His, besonders bevorzugt Val, Phe, Cys, Met, Trp, Asn, Arg oder His, ganz besonders bevorzugt Asn oder Arg;

5

x⁵ = eine Aminosäure ausgewählt aus der Gruppe Gly, Ser, Cys, Met, Trp, Asn, Glu, Lys oder Arg, besonders bevorzugt Gly, Ser, Cys, Met, Asn, Glu, Lys oder Arg, ganz besonders bevorzugt Gly oder Lys;

10

x⁶ = eine Aminosäure ausgewählt aus der Gruppe Phe, Pro, Ser, Cys, Trp, Tyr oder Gln, besonders bevorzugt Phe, Ser, Cys, oder Tyr, ganz besonders bevorzugt Cys.

- 15 Die Variablen x¹ bis x⁶ in der Sequenz His-x¹-His-x²-x³-x⁴-Cys-x⁵-x⁶-Cys können unabhängig voneinander die verschiedenen bevorzugten Bedeutungen haben, wobei einzelne Variablen bis maximal fünf der Variablen eine Aminosäure ausgewählt aus der Gruppe Gly, Ala, Val, Leu, Ile, Phe, Pro, Ser, Thr, Cys, Met, Trp, Tyr, Asn, Gln, 20 Asp, Glu, Lys, Arg, His bedeuten.

Besonders bevorzugte Peptidfragmente sind Fragmente mit den Sequenzen

25 His-Gln-His-Glu-Gly-Arg-Cys-Lys-Glu-Cys

His-Asn-His-Arg-Tyr-Gly-Cys-Gly-Cys-Cys

His-Arg-His-Gly-Thr-Asn-Cys-Leu-Lys-Cys

30

His-Ile-His-Gln-Ser-Asn-Cys-Gln-Val-Cys.

- Die oben genannte Proteinfragmentsequenz wird durch die erfindungsgemäßen Nukleinsäurefragmente kodiert. Hierbei ist der 35 degenerierte genetische Code zu berücksichtigen. Die erfindungsgemäßen Nukleinsäurefragmente können prinzipiell in beliebigen Nukleinsäuren enthalten sein. Vorteilhafterweise werden die Nukleinsäurefragmente in Vektoren so inseriert, daß die Herstellung von zusammengesetzten Nukleinsäuresequenzen (= Genkonstrukte), die für die erfindungsgemäßen Fusionsproteine 40 kodieren, ermöglicht wird. Diese Genkonstrukte können zur Expression vorteilhaft in einen geeigneten Wirtsorganismus verbracht werden, der eine optimale Expression des Fusionsproteins ermöglicht. Geeignete Vektoren sind dem Fachmann wohl bekannt 45 und können beispielsweise aus dem Buch Cloning Vectors (Eds. Pouwels P. H. et al. Elsevier, Amsterdam-New York-Oxford, 1985, ISBN 0 444 904018) entnommen werden. Unter Vektoren sind außer



Plasmiden auch alle anderen dem Fachmann bekannten Vektoren wie beispielsweise Phagen, Viren, Transposons, IS-Elemente, Plasmide, Cosmide, lineare oder zirkuläre DNA zu verstehen. Diese Vektoren können sich autonom im Wirtsorganismus replizieren oder chromosomal repliziert werden.

Unter den erfindungsgemäßen Nukleinsäuresequenzen sind Sequenzen zu verstehen, die mit einem oder mehreren Regulationssignalen vorteilhafterweise zur Erhöhung der Genexpression funktionell verknüpft wurden. Diese Sequenzen können 3' und/oder 5' Terminale regulatorische Sequenzen zur Steigerung der Expression sein, die je nach ausgewähltem Wirtorganismus und Gen für eine optimale Expression ausgewählt werden. Beispielsweise handelt es sich bei diesen regulatorischen Sequenzen um Sequenzen an die Induktoren oder Repressoren binden und so die Expression der Nukleinsäuren regulieren. Das Genkonstrukt kann außerdem vorteilhafterweise auch eine oder mehrere sogenannte "enhancer Sequenzen" funktionell verknüpft mit dem Promotor enthalten, die eine erhöhte Expression der Nukleinsäuresequenz ermöglichen. Dies kann beispielsweise über eine verbesserte Wechselwirkung zwischen RNA-Polymerase und DNA erfolgen. Auch am 3'-Ende der DNA-Sequenzen können zusätzliche vorteilhafte Sequenzen inseriert werden wie weitere regulatorische Elemente oder Terminatoren. Die erfindungsgemäßen Nukleinsäurefragmente werden vorteilhaft im Vektor so inseriert, daß sie die N-terminale Region des Fusionsproteins bilden. Sie können jedoch auch am C-Terminus lokalisiert sein oder aber innerhalb des Proteins liegen, wobei jedoch die Funktion des Proteins nicht beeinflußt sein darf und ein Ausschneiden aus dem Fusionsprotein nicht mehr möglich ist.

Die regulatorischen Sequenzen sollen die gezielte Expression der Gene und der Proteinexpression ermöglichen. Dies kann beispielsweise je nach Wirtsorganismus bedeuten, daß das Gen erst nach Induktion exprimiert oder überexprimiert wird, oder daß es sofort exprimiert und/oder überexprimiert wird.

Vorteilhafte Regulationssequenzen sind beispielsweise in Promotoren wie *cos*-, *tac*-, *trp*-, *tet*-, *trp-tet*-, *lpp*-, *lac*-, *lpp-lac*-, *lacIq*-, *T7*-, *T5*-, *T3*-, *gal*-, *trc*-, *ara*-, *SP6*-, λ -*P_R*- oder im λ -*P_L*-Promotor enthalten, die vorteilhafterweise in gram-negativen Bakterien Anwendung finden. Weitere vorteilhafte Regulationssequenzen sind beispielsweise in den gram-positiven Promotoren *amy* und *SPO2*, in den Hefe- oder Pilzpromotoren *ADC1*, *MFA*, *AC*, *P-60*, *CYC1*, *GAPDH*, *TEF*, *rp28*, *ADH* oder in den Pflanzenpromotoren *CaMV/35S*, *SSU*, *OCS*, *lib4*, *usp*, *STLS1*, *B33*, *nos* oder im Ubiquitin- oder Phaseolin-Promotor enthalten. In diesem Zusammenhang sind auch die Promotoren der Pyruvatdecarboxylase und der Methanol-

7 0 1 0 5 0 0

oxidase aus beispielsweise Hansenula vorteilhaft. Es können auch künstliche Promotoren für die Regulation verwendet werden.

Die regulatorischen Sequenzen bzw. Faktoren können dabei vorzugsweise die Genexpression des eingeführten Gens positiv beeinflussen und dadurch erhöhen. So kann eine Verstärkung der regulatorischen Elemente vorteilhafterweise auf der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie die genannten Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird.

Vorteilhafterweise enthalten die erfindungsgemäßen Nukleinsäuresequenzen auch Signale, die eine Sekretion der Proteine ins Medium oder in Zellkompartimente ermöglichen. Als Beispiel für derartige Sequenzen seien die typischen Signalsequenzen wie beispielsweise die Signalsequenz von ompA (E. coli-Membranprotein) genannt. Daneben können weitere vorteilhafte Sequenzen wie die Sequenz des α -Faktor enthalten sein oder sog. YACS (= Yeast artificial chromosomes) verwendet werden.

Die erfindungsgemäßen Genkonstrukte (= Nukleinsäuresequenzen) enthalten vorteilhafterweise außerdem noch Sequenzen, die eine Abspaltung der erfindungsgemäßen Proteinfragmente mit der oben genannten Sequenz vom N- oder C-Terminus des Fusionsproteins bevorzugt vom N-Terminus ermöglichen. Diese Sequenzen kodieren beispielsweise für Schnittstellen verschiedenster Proteasen wie beispielsweise für Faktor Xa, Enterokinase, humanen Renin, Carboxypeptidase A, Thrombin, Trypsin, Dipeptidylpeptidasen, Papain, Plasmin, Pepsin oder sonstigen Proteasen. Bevorzugt werden Schnittstellen für Faktor Xa, humanen Renin, Dipeptidylpeptidasen, Carboxypeptidase A oder Enterokinase, da diese Enzyme eine hohe Spezifität besitzen und so ein unerwünschter Verdau des zu reinigenden Proteins vermieden werden kann. Werden andere Proteasen verwandt, so ist darauf zu achten, daß keine Schnittstellen innerhalb des zu reinigenden Proteins sind. Das Proteinfragment kann auch über eine Bromcyanspaltung [z.B 2-(2-Nitrophenylsulfenyl)-3-brom-3'-methyldindolinium, Hydroxylamin etc.] in Gegenwart von Ameisensäure entfernt werden. Hierbei ist jedoch in der Regel eine Rückfaltung des Proteins erforderlich, was dazu führt, daß diese Methode weniger bevorzugt ist. Auch eine Abspaltung des Proteinfragments über einen Exoproteaseverdau (kinetisch kontrolliert) ist möglich. Hierbei entstehen jedoch in der Regel Produktgemische. Bevorzugt werden Schnittstellen verwendet, die eine Entfernung der Proteinfragmente ohne das Zurücklassen von Proteinfragmentresten im zu reinigenden Protein ermöglichen. Können im Fusionsprotein die erfindungsgemäßen Proteinfragmente

M 2 1 . 0 5 . 9 9
8

ohne Funktionsverlust und ohne sonstige Nachteile toleriert werden, so kann auf eine spezielle Stelle zur Abtrennung des Proteinfragments verzichtet werden.

- 5 Als Vektoren sind prinzipiell alle Vektoren geeignet, die eine Expression in pro- oder eukaryontischen Zellen ermöglichen. Dabei können Vektoren verwendet werden, die nur in einer Gattung replizieren oder solche, die in mehreren Gattungen replizieren (= sog. shuttle-vectoren). Vorteilhafte Vektoren sind beispielsweise Plasmide wie die E. coli-Plasmide pEGFP, pLG338, pACYC184, pBR322, pUC18, pUC19, pKC30, pRep4, pHS1, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III¹¹³-B1, λ gt11 oder pBdCI, bevorzugt pEGFP, in Streptomyces pIJ101, pIJ364, pIJ702 oder pIJ361, in Bacillus pUB110, pC194 oder pBD214, in Corynebacterium pSA77 oder pAJ667, in Pilzen pALS1, pIL2 oder pBB116, in Hefen 2 μ M, pAG-1, YE6, YE13 oder pEMBLYe23 oder in Pflanzen pLGV23, pGHIac⁺, pBIN19, pAK2004 oder pDH51. Die genannten Plasmide stellen eine kleine Auswahl der möglichen Plasmide dar. Weitere Plasmide sind dem Fachmann wohl bekannt und können beispielsweise aus dem oben genannten Buch Cloning Vectors (Eds. Pouwels P. H. et al. Elsevier, Amsterdam-New York-Oxford, 1985, ISBN 0 444 904018) entnommen werden.

- 25 In einer weiteren Ausgestaltungsform des Vektors kann die erfindungsgemäße Nukleinsäuresequenz auch vorteilhafterweise in Form einer linearen DNA in die Mikroorganismen eingeführt werden und über heterologe oder homologe Rekombination in das Genom des Wirtsorganismus integriert werden. Diese lineare DNA kann aus einem linearisierten Vektor wie einem Plasmid oder nur aus der Nukleinsäure, d.h. dem Nukleinsäurefragment und dem Gen für das Protein (= Fusionsproteingen) sowie gegebenenfalls weiteren regulatorischen Sequenzen bestehen.

- 35 Als Wirtsorganismen für das erfindungsgemäße Genkonstrukt kommen prinzipiell alle prokaryontischen oder eukaryontischen Organismen in Frage. Vorteilhafterweise werden als Wirtsorganismen Mikroorganismen wie gram-positive oder gram-negative Bakterien, Archaeobakterien, Pilze, Hefen, tierische oder pflanzliche Zellen wie Drosophila speziell D. melanogaster, Maus, Zebrafisch oder Tabak verwendet. Bevorzugt werden gram-positive oder gram-negative Bakterien, Pilze oder Hefen, besonders bevorzugt die Gattungen Escherichia, Bacillus, Streptomyces, Aspergillus oder Saccharomyces, ganz besonders bevorzugt E. coli verwendet.

- 45 Besonders bevorzugt sind folgende Kombinationen aus Vektor und Wirtsorganismus wie Escherichia coli und seinen Plasmiden und Phagen und den dazugehörigen Promotoren sowie Bacillus und seine

M 21.05.99

Plasmide und Promotoren, Streptomyces und seine Plasmide und Promotoren, Aspergillus und seine Plasmide und Promotoren oder Saccharomyces und seine Plasmide und Promotoren zu verstehen.

- 5 Die erfindungsgemäßen Fusionsproteine lassen sich wie oben beschrieben in einem Verfahren herstellen, wobei die erfindungsgemäßen Nukleinsäurefragmente, die für die Proteinfragmente mit der oben genannten Sequenz kodieren, mit einem Gen, das für die zu reinigenden Proteine kodiert, und gegebenenfalls weiteren
10 vorteilhaften Sequenzen wie Promotor- und/oder Enhancersequenzen, Schnittstellen für Proteasen etc. fusioniert werden. Dazu wird falls erforderlich eine geeignete Restriktionsenzymchnittstelle zwischen dem Nukleinsäurefragment und dem Gen des zu reinigenden Proteins eingeführt und dieses Konstrukt über geeignete Schnitt-
15 stellen in einen Vektor inseriert. Derartige Methoden sind dem Fachmann bekannt und können beispielsweise aus den Lehrbüchern von Sambrook, J. et al. (1989) Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, von F.M. Ausubel et al. (1994) Current protocols in molecular biology, John Wiley
20 and Sons oder D.M. Glover et al., DNA Cloning Vol.1, (1995), IRL Press (ISBN 019-963476-9) entnommen werden. Weitere vorteilhafte Vektoren sind die Pichia pastoris Vektoren pPic und pGap. Diese Hefe ist auch ein geeigneter Wirtsorganismus für die Protein-expression.

- 25 Die erfindungsgemäßen Proteinfragmente eignen sich zur Herstellung von Fusionsproteinen, die mit Hilfe der Proteinfragmente leicht, kostengünstig und effizient aufgereinigt werden können. Die erfindungsgemäßen Proteinfragmente und Fusionsproteine können
30 so vorteilhafterweise sehr selektiv in guten Ausbeuten, mit hoher Reinheit gereinigt werden. Die erfindungsgemäßen Proteinfragmente und damit die aus ihnen hergestellten Fusionsproteine zeichnen sich vorteilhaft durch eine im Vergleich zur Helicobacter pilori ATPase-439 Sequenz um mindestens 1,5 fach stärkere Bindung an die
35 immobilisierten Metallionen aus.

- Zur Herstellung der Fusionsproteine sind prinzipiell alle Proteine geeignet. Bevorzugt werden Proteine verwendet, die eine
40 biologische Wirkung in Mensch, Tier oder Pflanze zeigen oder die für organische Synthese interessant sind. Hierbei handelt es sich beispielsweise um Proteine wie Enzyme, Hormone, Speicher- oder Bindungs- oder Transportproteine. Beispielhaft seien Proteine wie Hydrolasen wie Lipasen, Esterasen, Amidasen, Nitrilasen, Prote-
45 asen, Mediatoren wie Cytokine z.B. Lymphokine wie MIF, MAF, TNF, Interleukine wie Interleukin 1, Interferone wie γ -Interferon, tPA, Hormon wie Proteohormone, Glykohormone, Oligo- oder Polypeptidhormone wie Vassopressin, Endorphine, Endostatin, Angiostatin,

10

Wachstumsfaktoren Erythropoietin, Transkriptionsfaktoren, Integrine wie GPIIb/IIIa oder $\alpha_v\beta_{III}$, Rezeptoren wie die verschiedenen Glutamatrezeptoren, Angiogenesefaktoren wie Angiotensin genannt.

5

Das erfindungsgemäße Verfahren zur Reinigung von Fusionsproteinen ermöglicht beispielsweise die Reinigung von Proteinen aus natürlichen Quellen wie Pflanzen- oder Tierextrakten, Pflanzen- oder Tierzell-Lysaten, aus Kulturmedien, Fermentationsbrühen oder aus

10 Synthesebrühen um nur einige beispielhaft zu nennen.

Das erfindungsgemäße Verfahren umfaßt folgende Reaktionsschritte:

15 a) Die Flüssigkeit(en), die das Fusionsprotein enthalten, werden so mit immobilisierten Metallionen in Kontakt gebracht, daß sich zwischen den Metallionen und dem Fusionsprotein eine Affinitätsbindung ausbilden kann,

20 b) Entfernen nicht gebundener in der Flüssigkeit enthaltener Substanzen,

c) Eluieren des gebundenen Fusionsproteins, in dem die Affinitätsbindung durch Änderung des Flüssigkeitsmediums aufgehoben wird und

25

d) Sammlung des gereinigten Fusionsprotein.

Vorteilhaft wird das Fusionsprotein in einem geeigneten Wirtsorganismus (siehe oben) vor der Reinigung exprimiert um die Ausbeute an Fusionsprotein zu erhöhen. Der Wirtsorganismus wird in einem geeigneten synthetischen oder komplexen Medium, das eine Kohlenstoffquelle, eine Stickstoffquelle und gegebenenfalls anorganische Salze, Vitamine und Spurenelemente enthält, bei geeigneter Temperatur und Belüftung angezogen.

35

Je nach dem ob das Fusionsprotein aus den Zellen exkretiert wird oder nicht, werden die Zellen zunächst aufgeschlossen und die Zellen oder Zelltrümmer vorteilhafterweise abgetrennt. Für den Zellaufschluß werden dem Fachmann bekannte Methoden verwendet wie

40 Ultraschall, French press, Enzymaufschluß, osmotischer Schock und andere mehr. Die Abtrennung der Zellen oder Zelltrümmer kann beispielsweise über Zentrifugation oder Filtration erfolgen. Eine Abtrennung der Zellen oder Zelltrümmer ist jedoch nicht zwingend erforderlich.

45

11

Die die Fusionsproteine enthaltende Flüssigkeit wird anschließend mit den immobilisierten Metallionen in Kontakt gebracht, so daß sich eine Affinitätsbindung zwischen dem Fusionsprotein und den Metallionen ausbilden kann. Die Bindung erfolgt bei pH-Werten größer 7, beispielsweise vorteilhaft bei pH 7,0 bis 9,0 bevorzugt zwischen pH 7,5 bis 8,0. Vorteilhafte Puffer sind einzelne Puffer oder Puffergemische wie beispielsweise 50 bis 1000 mM Puffer wie 50 mM Tris/HCl pH 8,0 + 150 mM NaCl, 100 mM NaAcetat pH 7,7 + 500 mM NaCl, 20 mM NaPhosphat pH 7,7 + 500 mM NaCl oder 50 mM Tris/HCl pH 8,0 + 150 mM NH_4Cl . Diese Puffer ermöglichen ein Auftragen der Fusionsproteine auf die immobilisierten Metallionen. Im einfachsten besonders vorteilhaften Fall werden die Fusionsproteine direkt im zum Aufschluß verwendeten Puffer oder im Inkubationsmedium mit den immobilisierten Metallionen in Kontakt gebracht. Vorteilhafterweise werden die Flüssigkeiten und die immobilisierten Metallionen in einer üblichen Chromatographiesäule miteinander in Kontakt gebracht. Dies erleichtert das Entfernen nicht gebundener Substanzen beispielsweise Proteine durch Waschen der Säule mit einem geeigneten Puffer. Geeignete Puffer sind Puffer, die die Bindung der erfindungsgemäßen Proteinfragmente oder der Fusionsproteine an die Metallionen nicht stören und Verunreinigungen entfernen können. Solche Puffer haben vorzugsweise einen pH-Wert größer pH 7, vorteilhaft einen pH-Wert zwischen pH 7,0 bis 9,0, bevorzugt zwischen pH 7,5 bis 8,0. Auch batch-Ansätze, bei denen die immobilisierten Metallionen in einem Gefäß vorgelegt werden und die Flüssigkeiten anschließend zugegesetzt werden oder umgekehrt, können so gereinigt werden. Für diese batch-Ansätze können die genannten Puffer verwendet werden. Zwischen den einzelnen Waschschritten werden die Ansätze vorteilhafterweise zentrifugiert oder filtriert.

Für die Immobilisierung der Metallionen eignen sich prinzipiell alle üblichen Trägermaterialien, die sich leicht derivatisieren lassen, keine oder nur geringe unspezifische Adsorption aufweisen, gute physikalische, mechanische und chemische Stabilität aufweisen und eine hohe äußere und innere Oberfläche aufweisen. Geeignete Materialien sind beispielsweise käuflich erwerbbar von Pharmacia LKB, Schweden (SephacroseTM 6B oder SuperoseTM), Pierce, USA (immobilisierte Iminodiessigsäure I oder II, immobilisiertes Tris(carboxymethyl)ethylendiamin), Sigma, USA (immobilisierte Iminodiessigsäureagarose), Boehringer Mannheim, Deutschland (Zinkchelatarose) oder Toyo Soda, Japan (TSKgel Chelate-5PW). In EP-B-0 253 303 werden weitere geeignete Materialien beschrieben. Weitere geeignete vorteilhafte Materialien sind Materialien wie Ni-beschichtete Mikrotiterplatten (Nickel-chelate coated

12

Flashplate®, New life science products) oder Magnetpartikel oder speziell mit Metallionen behandelte und bindende Membranen.

Die verschiedenen Metallionen werden in geeigneten Materialien vorteilhaft über Gruppen wie IDA (= Iminodiacessigsäure), NTA (= Nitrilotriessigsäure) oder TED (= Tris(carboxymethyl)ethylen-diamin gebunden. Geeignete Metallionen sind Co, Cu, Fe, Ca, Mg, Ni, Al, Cd, Hg oder Zn, bevorzugt werden Fe, Ni oder Cu, besonders bevorzugt Ni oder Cu, ganz besonders bevorzugt Ni. Die Beladung der Materialien mit Metallionen erfolgt vorteilhaft mit 0,1 bis 0,4 M Lösungen der Metallsalze in wäßriger, ungepufferter Lösung.

Nach dem Waschen wird das Fusionsprotein mit einem geeigneten Puffer eluiert. Dieser Puffer hebt die Affinitätsbindung zwischen dem Fusionsprotein und den immobilisierten Metallionen auf. Die Fusionsproteine können über einen pH-Gradienten (niedrige pH-Werte < pH 7,0 wirken eluierend), kompetitive Liganden wie Imidazol, organische Lösungsmittel wie Aceton oder Ethanol, chelatisierende Agentien wie EDTA oder NTA und/oder Detergentien wie Tween 80 eluiert werden. Bevorzugt wird eine Elution über kompetitive Liganden wie Imidazol und/oder Detergentien. Imidazol wird in einem Bereich von 0,05 bis 0,7 M, bevorzugt von 0,1 bis 0,5 M zur Elution verwendet. Die kompetitiven Liganden und/oder Detergentien werden vorteilhaft in einem Puffer eingesetzt, aber auch die Verwendung in Wasser ist möglich. Vorteilhafte Puffer sind Puffer, die den Puffern entsprechen, die für das Auftragen auf die immobilisierten Metallionen verwendet wurden. Dies hat den Vorteil, daß keine unerwünschten Wechselwirkungen zwischen dem Säulenmaterial, den gebundenen Proteinen und dem Puffer auftreten. Vorteilhafte Puffer haben vorzugsweise einen pH-Wert größer pH7, vorteilhaft einen pH-Wert zwischen pH 7,0 bis 9,0, bevorzugt zwischen pH 7,5 bis 8,0. Bevorzugt werden diese Puffer über einen steigenden Gradienten aufgebracht. Im Falle, daß mit einem pH-Gradient eluiert wird, können Puffer mit einem pH-Wert kleiner pH 7,0 und/oder Säuren verwendet werden. Das eluierte Fusionsprotein wird gesammelt und kann sofort verwendet werden, oder aber falls erforderlich und gewünscht weiter behandelt werden. Geeignete Auftrags- und Elutionspuffer sind beispielsweise dem Lehrbuch Protein Purification (Eds. J.C. Janson, L. Rydén, VCH Publisher Inc., 1989, Seite 227 bis 251) zu entnehmen.

Das erfindungsgemäße Proteinfragment kann mit den oben beschriebenen Methoden wie Bromcyan- oder Proteasespaltung entfernt werden. Hierbei können Reste des Proteinfragments im Molekül enthalten bleiben oder aber total aus dem zu reinigenden Protein

13

abgespalten werden. Vorteilhaft wird das Proteinfragment rückstandslos aus dem Protein entfernt.

Vorteilhaft geeignete Proteinfragmente für die Herstellung von
5 Fusionsproteinen lassen sich durch folgendes erfindungsgemäße
Verfahren screenen. Die Erfindung betrifft ein Verfahren zur
Herstellung von Proteinfragmenten, die in der Lage sind mit
immobilisierten Metallionen eine reversible Affinitätsbindung
einzugehen, dadurch gekennzeichnet, daß man folgende Schritte
10 durchführt:

- a) Herstellung einer Nukleinsäurebibliothek ausgehend von einer
beliebigen Nukleinsäuresequenz, die für ein Proteinfragment
der Sequenz

15

His-X¹-His-X²-X³-X⁴-Cys-X⁵-X⁶-Cys,

kodiert, wobei die Histidin- und Cysteinreste der Sequenz
in der Nukleinsäurebibliothek konserviert werden,

20

- b) Fusion der Nukleinsäuren der Bibliothek an ein Reportergen,
das die Detektion des durch die erhaltene Nukleinsäure
kodierten Fusionsproteins über seine Bindung an die
immobilisierten Metallionen ermöglicht und

25

- c) Auswahl der Nukleinsäuresequenzen, die eine gegenüber der
in der natürlichen *Helicobacter pilori* ATPase-439 Sequenz
um mindestens 1,5 fach stärkere reversible Bindung an die
immobilisierten Metallionen aufweisen.

30

Die Nukleinsäurebibliothek ausgehend von der oben genannten
Sequenz läßt sich über dem Fachmann bekannte Methoden zur
Mutagenese erstellen. Hierzu kann die Sequenz beispielsweise
einer "site directed mutagenesis" unterzogen werden wie sie
35 in D.M. Glover et al., DNA Cloning Vol.1, (1995), IRL Press
(ISBN 019-963476-9), Kapitel 6, Seite 193 ff beschrieben wird.

Von Spee et al. (Nucleic Acids Research, Vol. 21, No. 3, 1993:
777 - 778) wird eine PCR-Methode unter Verwendung von dITP zur
40 zufälligen Mutagenese beschrieben.

Die Verwendung einer "in vitro" Rekombinationstechnik für die
molekulare Evolution wird von Stemmer (Proc. Natl. Acad. Sci.
USA, Vol. 91, 1994: 10747 - 10751) beschrieben.

45

14

Von Moore et al. (Nature Biotechnology Vol. 14, 1996: 458 - 467) wird die Kombination der PCR- und Rekombinationsmethode beschrieben.

- 5 Die Verwendung einer "in vitro" Rekombinationstechnik für die molekulare Evolution wird von Stemmer (Proc. Natl. Acad. Sci. USA, Vol. 91, 1994: 10747 - 10751) beschrieben. Auch die Kombination der beiden Methoden kann verwendet werden.
- 10 Die Verwendung von Mutationsstämmen mit Defekten im DNA-Reparatursystem wird von Bornscheuer et al. (Strategies, 11, 1998: 16 - 17) beschrieben. Von Rellos et al. wird eine PCR-Methode unter Verwendung von nichtäquimolaren Nucleotidmengen (Protein Expression and Purification, 5, 1994: 270 - 277) beschrieben.
- 15

- 20 Vorteilhaft läßt sich die Nukleinsäurebibliothek über eine PCR-Technik mit Hilfe zweier komplementärer, degenerierter Oligonukleotide (= sog. wobble-primer), wie in den Beispielen beschrieben, erzeugen. Wichtig ist, daß die in der Sequenz enthaltenen Histidin- und Cysteinreste konserviert werden.

- Zum Screening auf Proteinfragmente mit einer verbesserten Metallbindungsaffinität wird das erfindungsgemäße Nukleinsäurefragment, das für das Proteinfragment kodiert, an ein Reportergen fusioniert. Vorteilhafte Reportergene ermöglichen eine leichte Detektion der Bindung an die immobilisierten Metallionen über beispielsweise eine Bindung von mit einem Fluoreszenzfarbstoff markierten Antikörpern, die gegen das Reportergen gerichtet sind, oder über selbst fluoreszierende Proteine wie das vorteilhafte bevorzugte egf-Protein von E. coli (= green fluorescent protein, siehe Prasher et al., Gene 111 (2), 1992: 229 - 233) oder das bevorzugte Biolumineszenzprotein von Aequoria victoria oder über Licht generierende Proteine wie das Luzeferin/Luzeferase-System. Besonders bevorzugt wird ein gfp-Proteinmutante (= egfp = enhanced green fluorescence protein) mit einer 35fachen höheren Fluoreszenzaktivität, die durch zwei Punktmutationen an Position 64 Austausch von Phe gegen Leu und Position 65 Austausch von Ser gegen Thr verursacht wird, verwendet. Diese Proteinmutante hat den Vorteil gegenüber dem Wildtypprotein, daß es löslich ist und keine sog. "Inclusion Bodies" bildet. Die Verwendung des egfp-Proteins ermöglicht die Lokalisierung und Quantifizierung der Proteinkonzentration in jeder Phase der Reinigung der Proteine ohne Eingriff in die Reinigung und ohne Verwendung von weiterer Kofaktoren oder Substrate (Poppenborg et al., J. Biotechnol., 45 58 (2), 1997, 77 - 88). Außerdem zeichnet sich das egfp-Protein durch eine hohe Stabilität gegenüber einem weiten pH-Bereich (pH 5,5 bis 12), Bleichung durch Photooxidation, oxidierende und

15

schwach reduzierende Agentien wie 2% Mercaptoethanol aus. Das Protein zeigt eine Abnahme der Fluoreszenz oberhalb 37 °C. Ebenfalls geeignet als Reportergen sind das gfp-uv- (blue fluorescence) und das eyfp- (yellow fluorescence) Protein.

- 5 Die Auswahl der geeigneten Sequenzen erfolgt im Vergleich zur Bindungsaffinität an die immobilisierten Metallionen der folgenden natürlichen *Helicobacter pilori* ATPase-439 Sequenz His-Ile-His-Asn-Leu-Asp-Cys-Pro-Asp-Cys. Die erfindungsgemäßen
- 10 Proteinfragmentsequenzen weisen eine um mindestens 1,5 fach stärkere reversible Bindung an die immobilisierten Metallionen auf, bevorzugt eine mindestens zweifach, besonders bevorzugt mindestens dreifach stärkere reversible Bindung auf. Vorteilhafte Sequenzen ermöglichen eine Proteinausbeute nach der Reinigung von
- 15 mindestens 20%, bevorzugt mindestens 30%, besonders bevorzugt von mindestens 40%, ganz besonders bevorzugt von mindestens 50%.
- 20 Das erfindungsgemäße Verfahren zum Screening der Nukleinsäurebibliothek eignet sich vorteilhaft für eine Automatisierung. Mit diesem Verfahren läßt sich einfach eine große Zahl von Nukleinsäurefragmenten bzw. Proteinfragmenten auf ihre Metallionenbindungsaffinität in einem sogenannten "High-Throughput-Screening" testen.
- 25 Mit den erfindungsgemäßen Proteinfragmenten lassen sich Proteine leicht detektieren. Im erfindungsgemäßen Verfahren zur Detektion von Proteinen werden aus einer Proteinmischung einzelne Proteine, die ein Proteinfragment mit den erfindungsgemäßen oben genannten Proteinfragmenten enthalten, über Antikörper nachgewiesen, die gegen das Proteinfragment gerichtet sind. Der Nachweis dieser
- 30 Fusionsproteine erfolgt vorteilhaft über mono- oder polyklonale Antikörper, die gegen das Proteinfragment (= Tag) gerichtet sind. Das Proteingemisch kann vorteilhaft vor dem Nachweis chromatographisch oder elektrophoretisch aufgetrennt werden und anschließend auf eine geeignete Membran (z.B. PVDF oder Nitrocellulose) mit
- 35 üblichen Methoden (siehe Sambrook et al.) überführt (= geblottet) werden. Anschließend wird diese Membran mit dem gegen das Tag gerichteten Antikörper inkubiert. Vorteilhaft wird die Membran mehrmals gewaschen und danach die gebundenen Antikörper über eine spezifische Reaktion mit einem beispielsweise Enzymkonjugierten
- 40 (z.B. alkalische Phosphatase, Peroxidase etc.) zweiten Antikörper der gegen die konstante Region des ersten gerichtet ist in einem sog. Western- oder Immuno-Blot nachgewiesen. Entsprechende Antikörper sind im Handel erhältlich. Im Falle der Verwendung von Magnetpartikel kann auf das Waschen verzichtet werden, die mit
- 45 Antikörpern beschichteten Magnetpartikel können über das Fischen mit Magneten gereinigt werden.

16 21 05 00

Die erfindungsgemäßen Proteinfragmente haben gegenüber den üblichen His-Tags bei der Proteindetektion den Vorteil, daß sie eine stärkere antigene Wirkung besitzen und so leichter zur Herstellung von Antikörpern gegen das Tag geeignet sind.

5

Die Erfindung wird durch die folgenden Beispiele weiter veranschaulicht.

Beispiele:

10

Für die Bindungstest an Metallchelatsäulen (= immobilisierte Metallionen) wurde die sog. "Chelating Sepharose Fast-Flow" von Pharmacia LKB, Uppsala, Schweden verwendet. Ampicillin, Imidazol, EDTA und alle weiteren Reagentien wurden von Fluka (Buchs,

15

Schweiz) bezogen. Das verwendete DNA Gel-Extraktions-Kit, das Midi Plasmid-Kit und das PrepSpin Plasmid Kit stammten von Qiagen (Hilden, Deutschland), die Restriktionsenzyme, die DNA-modifizierenden Enzyme, die T4-DNA-Ligase und die Taq-Polymerase stammten von MBI Fermentas (St. Leon-Rot, Deutschland). Der Taq

20

Dye Cycle Sequencing Kit wurde von Applied Biosystems (Weiterstadt, Deutschland) bezogen.

Für die Klonierungsversuche wurde der E. coli Stamm DH5 α (F⁻ endA1 hsdR17 [rk⁻, mk⁺] supE44 thi1 λ gyrA96 relA1 Δ (argF lacZya) U169) verwendet. Die Plasmide, die das Gen für das egf-Protein enthielten, wurden von Clontech USA bezogen. Die E. coli Stämme wurden bei 37 °C in Luria-Bertani Medium (= LB) mit 100 μ g/ml Ampicillin zur Selektion der Klone auf Transformation mit dem egfp-Vektor angezogen. Der Lysepuffer enthielt 50 mM NaPhosphatpuffer pH 8,0, 300 mM NaCl, 1 mg/ml Lysozym und 1 mM PMSF

30

(= Phenylmethansulfonylfluorid = spez. Trypsin- und Chymotrypsin-Inhibitor).

Die DNA-Methoden wie Ligationen, Restriktionen, PCR oder Transformationen etc. wurden wie bei Sambrook, J. et al. (1989) Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press oder F.M. Ausubel et al. (1994) Current protocols in molecular biology, John Wiley and Sons beschrieben durchgeführt. Für die Sequenzierung wurde die Fluoreszenzmarkierte Dideoxy-DNA-Sequenzierungsmethode verwendet. Die DNA-Sequenzierung wurde mit dem Taq Dye DeoxyTM Cycle Sequencing Kit (Applied Biosystems) und dem 373A DNA Sequencing System (Applied Biosystems) nach Angaben des Herstellers durchgeführt.

45

17

Beispiel 1: Herstellung zufällig mutagenisierter N-terminaler Metallbindungsstellen, die an das egfp-Gen gebundenen wurden, und des his6-egfp

5 Für die PCR-Reaktion wurde das Plasmid egfp und die zwei folgenden komplementäre Oligonukleotide

5'-GCAATACCATGGGGCATNNNCATNNNNNNNNNTGTNNNNNNNTGTGTGAGGAAGGGCGAG-3'

10 5'-CAGTTGGAATTCTAGAG-3'

verwendet. Im Falle des his6-egfp wurden die folgenden zwei komplementären Primer

15 5'-GCAATACCATGGGGCATCATCATCATCATGTGAGGAAGGGCGAG-3'

5'-CAGTTGGAATTCTAGAG-3'

verwendet.

20 Die für die PCR-Reaktionen verwendeten Bedingungen waren wie folgt:

Ansatz:	8	µl dNTP Mix (200 µmol)
25	10	µl 10 x ThermoPolPuffer (New England Biolabs)
	1	µl (100 pmol) Primer1
	1	µl (100 pmol) Primer2
	1	µl (100 ng) egfp Plasmid
	79,5	µl Wasser
30	1	µl Deep Vent Polymerase (New England Biolabs)

PCR-Programm	95 °C 7min
	95 °C 1min
	56 °C 1min 30x
35	72 °C 3min
	72 °C 7min

Die PCR-Produkte wurden jeweils mit NcoI und NotI verdaut und in den egfp-Vektor, der mit den gleichen Enzymen verdaut worden war, um Mutationen im Vektor auszuschließen legiert (siehe Figur 1).

40 Die PCR-Vektorligationen wurden zur Transformation von E. coli verwendet. Die Transformanden wurden auf LB-Agar mit 100 µg/ml Ampicillin plattiert und bei 37 °C inkubiert.

M 21.15.00
18

Beispiel 2: Kultivierungsbedingungen und Herstellung der Zell-Lysate

Transformierte Kolonien, die Fluoreszenz und einige die keine
5 Fluoreszenz zeigten, wurden ausgewählt und in 50 ml LB-Medium,
das 100 µg/ml Ampicillin enthielt, angezogen. Für das High-
Throughput-Screening wurden Kolonien ausgewählt, die Fluoreszenz
zeigten, und in sterilen Mikrotiterplatten, die 250 µl LB-Medium
mit 100 µg/ml Ampicillin enthielten, inkubiert. Nach Inkubation
10 wurden die Kulturen zentrifugiert. Die Pellets wurden in 2 ml
Lysepuffer resuspendiert, 20 Minuten auf Eis inkubiert und an-
schließend mit Ultraschall aufgeschlossen (zweimal, 5 Minuten mit
einem Branson Sonifier 250). Nach Zentrifugation (15 min, 4 °C,
20000 x g) wurden die verschiedenen egfp-Mutanten im Überstand
15 erhalten. Alle ausgewählten Klone wurden sequenziert (siehe Ta-
belle I und II). Klone, die keine Fluoreszenz zeigten, enthielten
Stopcodons in der Sequenz, so daß keine funktionellen Proteine
exprimiert wurden (siehe Tabelle I, A8, A13, M16a, Z4, Z11 und
Z13). In allen Versuchen wurde zur Quantifizierung der gebundenen
20 Proteine eine Fluoreszenzmessung durchgeführt, die in einem
weiten Bereich mit der gfp-Konzentration korrelierte (siehe Fi-
gur 2).

Beispiel 3: Low-Throughput-Screening mit Ni-NTA-Säulen von
Quiagen

25 600 µl der lysierten Zellen wurden auf eine Ni-NTA-Säule aufge-
bracht, zweimal mit 600 µl einer Waschlösung (50 mM NaPhosphat-
puffer, pH 8,0, 250 mM NaCl) gewaschen und anschließend mit einer
30 0,7 M Imidazollösung eluiert.

Beispiel 4: High-Throughput-Screening mit Membranfilterplatten

Für das High-Throughput-Screening wurden Membranfilter-Mikro-
titerplatten (MultiScreen 5 µm; Fa. Millipore, Molsheim, Deutsch-
35 land) verwendet. In jedes "Well" der Membranfilterplatten wurden
je dreimal 250 µl einer gerührten Chelatsepharosesuspension
gegeben. Nach jeder Zugabe wurde die Sepharose abzentrifugiert
(2 min, 23 °C, 350 rpm). Alle weiteren Schritte wurden in einem
Beckmann Biomek 2000 Roboter durchgeführt. Die Minisäulen in den
40 Wells wurden zweimal mit 250 µl Wasser gewaschen. Anschließend
wurde die Sepharose mit 250 µl einer Metallsalzlösung beladen
und dreimal mit 200 µl Puffer (50 mM NaPhosphat, pH 8,0, 250 mM
NaCl) äquilibriert. Für die verschiedenen Ansätze (Beladung mit
Metallionen) wurden je 0,3 M NiCl₂-, 0,3 M CuSO₄- oder 0,3 M
45 ZnCl₂-Lösungen verwendet. Es wurden wäßrige, ungepufferte Metall-
salzlösungen verwendet. 200 µl der Zell-Lysatüberstände wurden
auf jede dieser Minisäulen gegeben. Die Säulen wurden anschlie-

M 01.05.99
19

5 Bend zweimal mit 250 µl Äquilibrierungspuffer gewaschen. Die gebundenen Proteine wurden mit zweimal 100 µl 0,5 M Imidazol in Äquilibrierungspuffer eluiert. Die Chelatsepharose in den Filtern kann mit 250 µl 50 mM EDTA, 1 M NaCl in Wasser regeneriert werden und für weitere Screeningversuche verwendet werden.

Beispiel 5: IMAC-Versuche

10 Für den Versuch wurde ein üblicher Chromatographieaufbau aus einer Glassäule, zwei peristaltischen Pumpen zum Aufbringen der Lösungen, einem UV-Detektor (LKB UV-MII), einem Drucker (LKB RIC 102) und einem Fraktionssammler (LKB FRAC-200) gewählt. Alle Geräte stammten von der Firma Pharmacia. Die Säule wurde mit chelatisierender Sepharose Fast-Flow Gel (Pharmacia) gefüllt, mit 15 7 Bettvolumen deionisiertem Wasser gewaschen und mit 7 Bettvolumen 0,3 M NiCl_2 -Lösung mit den Metallionen beladen. Anschließend wurde die Säule mit 7 Bettvolumen IMAC-Puffer (50 mM NaPhosphat, pH 8,0, 250 mM NaCl) gewaschen und äquilibriert. 1 ml Proben der Zell-Lysate wurden auf die Säule mit einer Flußrate von 20 1,5 ml/min aufgebracht und mit 10 Bettvolumen IMAC-Puffer gewaschen. Die Elution der gebundenen Proteine erfolgt über einen ansteigenden Gradienten mit 0,5% einer 0,5 M Imidazollösung pro ml Elutionslösung und schließlich 5 Bettvolumen einer 0,5 M Imidazol/Wasserlösung. Die Proteinfraktionen wurden über UV- 25 Detektion ermittelt und gesammelt. Nach Abschluß wurde die Säule mit 50 mM EDTA/1 M NaCl-Lösung gewaschen und regeneriert. Dieser Waschschrift löst das Metall, gebundene Zellreste und Proteine von der Säule. Die eluierten Fraktionen wurden sowohl optisch als auch spektroskopisch untersucht. Einige der untersuchten 30 Klone zeigten keine Affinität zur Matrix (siehe Tab.I, M15, M16) während andere eine gute Bindung aufwiesen (siehe Tab.I, M13, Z5 und Z7). Die Klone M13 und Z5 eluierten in einer scharfen Bande von der Säule, während der Klon Z7 auf der Säule schmierte.

35 Beispiel 6: Vergleichsversuch mit egfp-Wildtyp und his-tag

Der Klon M13 wurde in einem Vergleichsversuch mit dem egfp-Wild-
typprotein und den üblichen his-tags verglichen. Das egfp-Wild-
typprotein bindet nicht an die Metallchelatsäulen. Nach dem
Waschen der Säule war keine Fluoreszenz mehr auf der Säule
40 detektierbar. Der Klon M13 bindet in einer scharfen Bande auf der Säule während die his-tag-Proteine über die gesamte Säule binden. Dies ist auf eine geringere Affinität zurückzuführen, die letztlich zu einer geringeren Kapazität der Säule führt. Die Proteinausbeute liegt im Falle von M13 mit 56% höher als
45 mit den his-tags mit 48%.

20

Tabelle I: Bindungsversuche an Ni-Metallchelatsäulen

	Klon	Aminosäuresequenz										
5	A6	His	Gln	His	Glu	Gly	Arg	Cys	Lys	Glu	Cys	gfp
	A8	His	Cys	His	Pro	Glu	Leu	Cys	Stop	Leu	Cys	gfp
	A13	His	Leu	His	Ser	Ile	Gly	Cys	Pro	Stop	Cys	gfp
	M13	His	Asn	His	Arg	Tyr	Gly	Cys	Gly	Cys	Cys	gfp
10	M14	His	Ser	His	Ser	Val	Gly	Cys	Phe	Phe	Cys	gfp
	M15	His	Gly	His	Thr	Leu	Ser	Cys	Gly	Leu	Cys	gfp
	M16	His	Ser	His	Thr	Leu	Arg	Cys	Lys	Gly	Cys	gfp
	M16a	His	Ser	His	Stop	Leu	Arg	Cys	Lys	Gly	Cys	gfp
15	Z4	His	Stop	His	Asn	Stop	Val	Cys	Ala	Thr	Cys	gfp
	Z5	His	Arg	His	Gly	Thr	Asn	Cys	Leu	Lys	Cys	gfp
	Z7	His	Ile	His	Gln	Ser	Asn	Cys	Gln	Val	Cys	gfp
	Z11	His	Thr	His	Ala	Ser	Gly	Cys	Stop	Stop	Cys	gfp
	Z13	His	Cys	His	Thr	Trp	Cys	Cys	Asn	Stop	Cys	gfp

20

Die Klone A6, A10, M13, Z5 und Z7 binden gut an die Metallchelatsäule, während die Klone M14, M15 und M16 keine Bindung zeigten.

Tabelle II: Weitere sequenzierte im High-Throughput-Screening durch Fluoreszenz aufgefallene Klone

25

	Klon	Aminosäuresequenz										
30	A1	His	Gly	His	Met	Glu	Arg	Cys	Leu	Val	Cys	gfp
	A2	His	Lys	His	Ala	Arg	Ser	Cys	Met	Gly	Cys	gfp
	A3	His	Phe	His	Thr	Val	Phe	Cys	Phe	Ser	Cys	gfp
	A4	His	Arg	His	Arg	Gly	Met	Cys	Thr	Ala	Cys	gfp
35	A12	His	Asp	His	Arg	Gly	Val	Cys	Gly	Leu	Cys	gfp
	A14	His	Asp	His	Glu	Arg	Leu	Cys	His	Asn	Cys	gfp
	X8	His	Gly	His	Gly	Asn	Arg	Cys	Cys	Gly	Cys	gfp
	X9	His	Arg	His	Gly	Thr	Ala	Cys	Met	Asp	Cys	gfp
40	X11	His	Ile	His	Ile	Met	Thr	Cys	Leu	Ser	Cys	gfp
	X12	His	Thr	His	Pro	Arg	Ser	Cys	Ala	Glu	Cys	gfp
	X15	His	Gly	His	Asp	Arg	Thr	Cys	Arg	Gly	Cys	gfp
	X16	His	Arg	His	Ala	Ile	Ser	Cys	Ile	Gly	Cys	gfp
45	X17	His	Ile	His	Arg	Gly	Asp	Cys	Tyr	Glu	Cys	gfp
	X18	His	His	His	Gly	Ser	Thr	Cys	Pro	Thr	Cys	gfp
	X19	His	His	His	Phe	His	Ser	Cys	Phe	Tyr	Cys	gfp
	Z8	His	Lys	His	Val	Asp	His	Cys	Gly	Arg	Cys	gfp

M 21.05.99
21

Klon	Aminosäuresequenz										
Z9	His	Ser	His	Leu	Thr	Leu	Cys	Leu	Gly	Cys	gfp
Z10	His	Thr	His	Gln	Ser	Gln	Cys	Gly	Arg	Cys	gfp
Z14	His	Arg	His	Leu	Phe	Trp	Cys	Ser	Glu	Cys	gfp

Beispiel 7: Metallbindungsaffinität

Viele der Klone zeigten eine bevorzugte Bindung an Ni^{2+} oder Cu^{2+} .
10 Im Falle von M13 konnte keine Bindung an Zn^{2+} beobachtet werden.
Bei Verwendung von Ni-Chelatsäulen zeigt der Klon M13 eine deutlich bessere Reinigung der Proteine im Vergleich zu den his-tags. Diese führten umgekehrt bei Verwendung von Cu-Chelatsäulen zu einem reineren Produkt im Vergleich zu M13, wobei jedoch, da die
15 Bindung an das Säulenmaterial in beiden Fällen sehr stark ist, durch die drastischen Elutionsbedingungen Cu-Ionen ausgewaschen wurden. Dies führt zu Produktverunreinigungen.

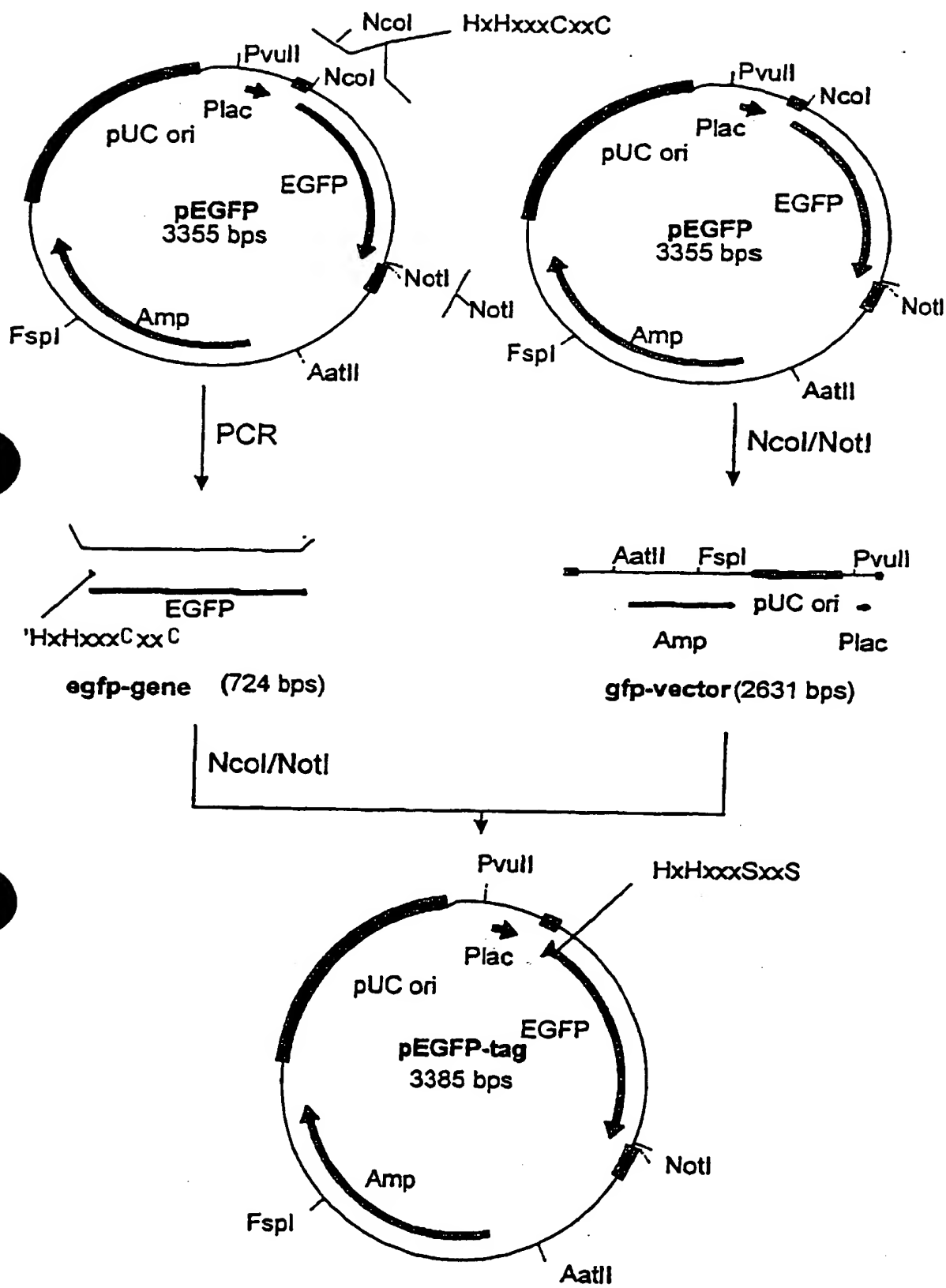
Beispiel 8: Vergleichsversuch zwischen Sequenz der ATPase-439
und erfindungsgemäßen Proteinfragmenten

Der Vergleichsklon der ATPase-439 wurde analog Beispiel 1 und 6 durchgeführt. Als Primer wurde folgender Primer
5'-GCAATACCATGGGGCATATTCATAATCTTGATTGTCCTGATTGT-3' verwendet. Die
25 anderen Primer und die PCR-Bedingungen waren wie unter Beispiel 1 beschrieben.

Der Versuch wurde mit dem Quiagen Ni-NTA Spin Kit unter nativen Bedingungen durchgeführt, unter diesen Bedingungen wurde wie beschrieben lysiert; die bereits fertig beladenen Säulen wurden mit
30 600 µl 50 mM Na-Phosphatpuffer, pH 8,0, 300 mM NaCl äquilibriert und 2 min bei 2000 rpm (= 420 x g) zentrifugiert. Anschließend wurden 600 µl des Lysats aufgebracht und wie 2 min zentrifugiert. Mit je 600 µl 50 mM Na-Phosphatpuffer, pH 8,0, 300 mM NaCl wurde
35 zweimal gewaschen. Danach wurde mit 600 µl 50 mM Na-Phosphatpuffer, pH 8,0, 300 mM NaCl, 0,5 M Imidazol wurde zweimal eluiert. Die Ausbeute an reinem Protein war mit dem Klon M13 1,5fach höher als der Metallbindungstelle des Vergleichsklons der ATPase-439. Klon M13 bindet damit besser und läßt sich auch besser eluieren.

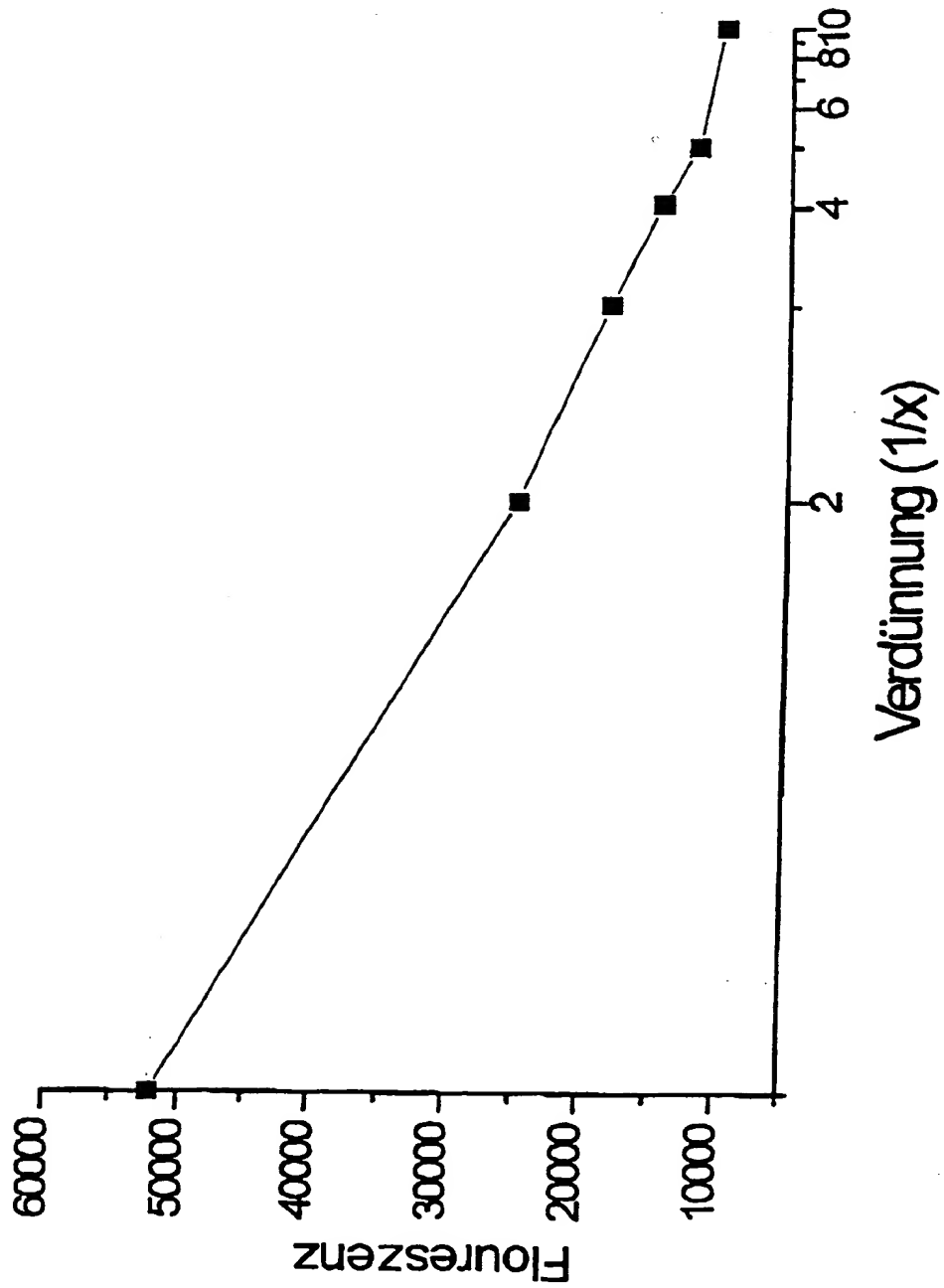
Figur 1

14.21.05.99



M 21.05.99

Figur 2



M 21.05.99

Neue Peptidfragmente zur Reinigung von Proteinen

Zusammenfassung

5

Die Erfindung betrifft neue Peptidfragmente, Fusionsproteine enthaltend die Peptidfragmente, Verfahren zur ihrer Herstellung sowie die Verwendung der Peptidfragmente. Die Erfindung betrifft auch ein Verfahren zur Reinigung der Fusionsproteine und ein

10 Verfahren zur Detektion von Proteinen.

Weiterhin betrifft die Erfindung Nukleinsäuren, die für die Peptidfragmente oder für die Fusionsproteine kodieren und Vektoren, die die Nukleinsäuren enthalten.

15

20

25

30

35

40

45